UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/992,987	11/14/2001	David W. Camarillo	41698-1045	7662

7590 03/04/2005

EXAMINER	
WONG, LESLIE	

ART UNIT	PAPER NUMBER
2167	

DATE MAILED: 03/04/2005

Alex L. Yip
Kaye Scholer LLP
425 Park Avenue
New York, NY 10022

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/992,987	CAMARILLO, DAVID W.	
	Examiner	Art Unit	
	Leslie Wong	2167	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *25 June 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-21* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-21* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *24 February 2002* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☒ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Priority

1. The Applicant's claim to domestic priority under 35 U.S.C. §119(e), based on a provisional of application serial number 60/314,722, filed on August 24, 2001, is acknowledged.

Claim Rejections - 35 USC § 101

2. Claims 4, 7-10, 11-14, 15-16 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The language of the claim raises a question as to whether the claim is directed merely to an abstract idea that is not tied to a technological art, environment or machine which would result in a practical application producing a concrete, useful, and tangible result to form the basis of statutory subject matter under 35 U.S.C 101.

Claims 4, 7-10, and 11-14 are system claims, however, there is no limitation containing hardware components such as a memory or processor. As a result, these claims are not statutory.

In order to overcome the 101 rejections on claims 15-16, it is suggested that the preamble of claim 15 be amended to recite the following:

"Software recorded in a computer-readable medium, said software including computer-readable instructions for performing a process for building a database, to store records corresponding to a plurality of data items, the process comprising:"

To expedite a complete examination of the instant application the claims rejected under 35 U.S.C. 101 (nonstatutory) above are further rejected as set forth below in anticipation of applicant amending these claims to place them within the four categories of invention.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 1-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Kane et al.** ("Kane") (U.S. Patent 6,389,429) in view of **Seilhamer et al.** ("Seilhamer") (U.S. Patent 6,023,659).

Regarding claim 1, **Kane** teaches a method for maintaining records in a database comprising:

a). **'receiving at least a collection of first data items and a collection of second data items'** as processing source databases because record layouts between various source databases 101, 102, and 103 and the target database will not be identical. It is likely that different databases contain information about the same audience member, but the data within the databases may be inconsistent (col. col. 7, lines 17-23; Fig. 1; col. 4, lines 30-49);

b). **'disposing the first data items in a plurality of fields arranged in a predetermined format to form a first assemblage'** as it is necessary to extract data from *each* source database so that the extracted data will be broken down into fields formatted to match the format of corresponding fields of the target database (col. 7, lines 17-23);

c). **'disposing the second data items in a plurality of fields arranged in the predetermined format to form a second assemblage'** as it is necessary to extract data from *each* source database so that the extracted data will be broken down into fields formatted to match the format of corresponding fields of the target database. Moreover, it may be desirable to use multiple source databases to update the target database (col. 7, lines 17-23; col. 8, lines 23-35);

e). **'maintaining a record, in the database, having a plurality of fields arranged in the predetermined format'** as the target database 108 of audience member records may have any desired record layout, but the indicated layout is suitable (col. 3, line 50 – col. 4, line 14); and

f). **'determining whether a first particular data item in the predetermined nomenclature in a selected field of the first assemblage is identical to a second particular data item in the predetermined nomenclature in a field of the second assemblage corresponding to the selected field'** as all source database records are first extracted, then, on a record by record basis, a matching record in the target database is identified and the data for the

matched record in the target database is updated (Fig. 3 and col. 13, lines 47-52)(col. 9, lines 24-52). A representative criteria for determining whether a match exists between a source and a target record - **a negative number is a miss**, 0 is used when no data for a field is source, and **a 100 represents a perfect match** (i.e., identical) (col. 10, lines 18-49).

g). **'if it is determined that the first particular data item in the predetermined nomenclature is identical to the second particular data item in the predetermined nomenclature, including in a field in the record a selected one of (1) the first data items in the predetermined nomenclature and (2) the second data items in the predetermined nomenclature'** as all source database records are first extracted, then, on a record by record basis, a matching record in the target database is identified and the data for the matched record in the target database is updated (Fig. 3 and col. 13, lines 47-52)(col. 9, lines 24-52). For each field group in the new target database record, the source database is specified for the fields name datasource_id, address_datasource_id, phone_datasource_id, and email_datasource_id, to indicate that the new target record information came from the particular source database (col. 12, lines 20-27).

d). **Kane** does not explicitly teach modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature.

Seilhamer, however, teaches **'modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature'** as a species filter may be employed to standardize all references to a particular species. For example, GenBank references

to "man, "homo sapeins", and "human" etc. as well as misspellings of these words are all converted to a standard term such as "human" (col. 27, lines 55-59).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to combine the teachings of the cited references because **Seilhamer's** teaching would have allowed **Kane's** to provide more accurate search results by standardizes the data from various sources to conform with the pre-determined nomenclature as suggested in (col. 27, lines 14-23; col. 27, lines 61-62).

Regarding claim 2, **Kane** further teaches **'wherein at (b) and (c) the first data items and the second data items are separated, rearranged, or combined to form the first assemblage and the second assemblage respectively'** as the source data is broken down into fields formatted to match the format of corresponding fields of the target database (col. 7, lines 17-23).

Regarding claim 3, **Kane** further teaches **'wherein at (f) the first data items and the second data items are assigned accuracy rankings and are selected based upon their accuracy rankings'** as the scoring criteria is applied to determine minimum threshold for a match and which record has the highest ranking for a match. Data in the target database always being kept up to date and only source records that have a higher accuracy ranking than the fields in the target database are updated (col. 9, lines 18-58 and abstract; col. 12, lines 60-67).

Regarding claim 4, **Kane** further teaches a system for maintaining records in a database comprising:

a). **'communications interface for receiving at least a collection of first data items and a collection of second data items'** as processing source databases because record layouts between various source databases 101, 102, and 103 and the target database will not be identical. It is likely that different databases contain information about the same audience member, but the data within the databases may be inconsistent (col. col. 7, lines 17-23; Fig. 1; col. 4, lines 30-49);

b). **'a converter for disposing the first data items in a plurality of fields arranged in a predetermined format to form a first assemblage, and for disposing the second data items in a plurality of fields arranged in the predetermined format to form a second assemblage'** as it is necessary to extract data from each source database so that the extracted data will be broken down into fields formatted to match the format of corresponding fields of the target database and when the data table from the extraction process from the source data is created, it will be known how fields in the table correspond, or map, to data in the target database per step 202 of Fig. 2. Moreover, it may be desirable to use multiple source databases to update the target database (col. 7, lines 17-23; col. 8, lines 23-35);

d). **'a database for maintaining a record, in the database, having a plurality of fields arranged in the predetermined format'** as the target database 108 of audience member

records may have any desired record layout, but the indicated layout is suitable (col. 3, line 50 – col. 4, line 14); and

e). **'a data masher for determining whether a first particular data item in the predetermined nomenclature in a selected field of the first assemblage is identical to a second particular data item in the predetermined nomenclature in a field of the second assemblage corresponding to the selected field, and for including in a field in the record a selected one of (1) the first data items in the predetermined nomenclature and (2) the second data items in the predetermined nomenclature, if it is determined that the first particular data item in the predetermined nomenclature is identical to the second particular data item in the predetermined nomenclature'** as all source database records are first extracted, then, on a record by record basis, a matching record in the target database is identified and the data for the matched record in the target database is updated (Fig. 3 and col. 13, lines 47-52)(col. 9, lines 24-52). A representative criteria for determining whether a match exists between a source and a target record - **a negative number is a miss**, 0 is used when no data for a field is source, and **a 100 represents a perfect match** (i.e., identical) (col. 10, lines 18-49). For each field group in the new target database record, the source database is specified for the fields name datasource_id, address_datasource_id, phone_datasource_id, and email_datasource_id, to indicate that the new target record information came from the particular source database (col. 12, lines 20-27).

c). **Kane** does not explicitly teach a normalizer for modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature.

Seilhamer, however, teaches **'modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature'** as a species filter may be employed to standardize all references to a particular species. For example, GenBank references to "man, "homo sapeins", and "human" etc. as well as misspellings of these words are all converted to a standard term such as "human" (col. 27, lines 55-59).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to combine the teachings of the cited references because **Seilhamer's** teaching would have allowed **Kane's** to provide more accurate search results by standardizes the data from various sources to conform with the pre-determined nomenclature as suggested in (col. 27, lines 14-23; col. 27, lines 61-62).

Regarding claim 5, **Kane** further teaches wherein **'each of the of the first assemblage of data items and the second assemblage of data items in the predetermined nomenclature are associated with data accuracy rankings'** as the scoring criteria is applied to determine minimum threshold for a match and which record has the highest ranking for a match (col. 9, lines 18-58 and abstract; col. 12, lines 60-67).

Regarding claim 6, **Kane** further teaches wherein **'the data masher selects the first data items and the second data items based upon the data accuracy rankings'** as source records

that have a higher accuracy ranking than the fields in the target database are updated (col. 9, lines 18-58 and abstract; col. 12, lines 60-67).

Regarding claim 7, **Kane** further teaches a system for maintaining records in a database comprising:

a). **'a converter for converting at least a first collection of data items and a second collection of data items to at least a first assemblage and second assemblage, respectively, the first and second assemblage each having data items in fields which are arranged in a predetermined format'** as record layouts between various source databases 101, 102, and 103 and the target database will not be identical and it is likely that different databases contain information about the same audience member, but that data within the databases may be inconsistent. It is necessary to extract data from *each* source database so that the extracted data will be broken down into fields formatted to match the format of corresponding fields of the target database and compare the source records and update the target database with selected source records that have a higher accuracy ranking (col. 7, lines 17-23; Fig. 1; col. 4, lines 30-49; col. 9, lines 18-57);

c). **'a database for maintaining a record having a plurality of fields arranged in the predetermined format'** as the target database 108 of audience member records may have any desired record layout, but the indicated layout is suitable (col. 3, line 50 – col. 4, line 14); and

d). **'a data masher for determining a value representing a number of corresponding fields in the first assemblage and the second assemblage having identical data items therein, the data masher based on the value selecting at least one of the data items in the first assemblage and the second assemblage to form the record'** as comparing the source records and update the target database with selected source records that have a higher accuracy ranking (col. 9, lines 18-57; col. 4, lines 30-50). All source database records are first extracted, then, on a record by record basis, a matching record in the target database is identified and the data for the matched record in the target database is updated (Fig. 3 and col. 13, lines 47-52)(col. 9, lines 24-52). A representative criteria for determining whether a match exists between a source and a target record - **a negative number is a miss**, 0 is used when no data for a field is source, and **a 100 represents a perfect match** (i.e., identical) (col. 10, lines 18-49);

b). **Kane** does not explicitly teach a normalizer for converting selected ones of the first assemblage of data items and the second assemblage of data items to conform to a predetermined nomenclature.

Seilhamer, however, teaches **'converting selected ones of the first assemblage of data items and the second assemblage of data items to conform to a predetermined nomenclature'** as a species filter may be employed to standardize all references to a particular species. For example, GenBank references to "man, "homo sapeins", and "human" etc. as well as misspellings of these words are all converted to a standard term such as "human" (col. 27, lines 55-59).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to combine the teachings of the cited references because **Seilhamer's** teaching would have allowed **Kane's** to provide more accurate search results by standardizes the data from various sources to conform with the pre-determined nomenclature as suggested in (col. 27, lines 14-23; col. 27, lines 61-62).

Regarding claim 8, **Kane** further teaches wherein **'the converter separates, rearranges, or combines data to form the first assemblage of data items and the second assemblage of data items'** as the source data is broken down into fields formatted to match the format of corresponding fields of the target database (col. 7, lines 17-23).

Regarding claim 9, Kane further teaches wherein **'each of the of the first assemblage of data items and the second assemblage of data items in the predetermined nomenclature are associated with data accuracy rankings'** as the scoring criteria is applied to determine minimum threshold for a match and which record has the highest ranking for a match (col. 9, lines 18-58 and abstract; col. 12, lines 60-67).

Regarding claim 10, **Kane** further teaches wherein **'the data masher selects data items in the first assemblage and the second assemblage based upon the data accuracy ratings'** as source records that have a higher accuracy ranking than the fields in the target database are updated (col. 9, lines 18-58 and abstract; col. 12, lines 60-67).

Regarding claim 11, **Kane** further teaches a system for maintaining records in a database comprising:

a). **'a converter for converting first records containing data items to uniform records having the data items organized in a uniform format, wherein the uniform records have one or more of the data items conforming to a predetermined nomenclature and a balance of the one or more of the data items not conforming to the predetermined nomenclature'** as record layouts between various source databases 101, 102, and 103 and the target database will not be identical and it is likely that different databases contain information about the same audience member, but that data within the databases may be inconsistent. It is necessary to extract data from *each* source database so that the extracted data will be broken down into fields formatted to match the format of corresponding fields of the target database and compare the source records and update the target database with selected source records that have a higher accuracy ranking (col. 7, lines 17-23; Fig. 1; col. 4, lines 30-49; col. 9, lines 18-57);

c). **'a data masher for selecting, from among the second records, at least a first selected record and a second selected record based on a value representing a number of corresponding fields in the first selected record and the second selected recorded having identical data items therein, and for selecting data items from the first selected record and second selected record to form a third data record, the third data record being stored in the database'** as comparing the source records and update the target database with selected source records that have a higher accuracy ranking (col. 9, lines 18-57; col. 4, lines 30-50). as all source database records are first extracted, then, on a record by record basis, a matching record

in the target database is identified and the data for the matched record in the target database is updated (Fig. 3 and col. 13, lines 47-52)(col. 9, lines 24-52). A representative criteria for determining whether a match exists between a source and a target record - **a negative number is a miss**, 0 is used when no data for a field is source, and **a 100 represents a perfect match** (i.e., identical) (col. 10, lines 18-49).

b). **Kane** does not explicitly teach a normalizer for converting the balance of the one or more data items to the predetermined nomenclature, producing a collection of second records each having the one or more of the data items that conform to a predetermined nomenclature and the balance of the one or more data items that are in the predetermined nomenclature.

Seilhamer, however, teaches '**normalizer for converting the balance of the one or more data items to the predetermined nomenclature, and producing a collection of second records each having selected data items organized in fields according to the uniform format, the selected data items conforming to the predetermined nomenclature**' as a species filter may be employed to standardize all references to a particular species. For example, GenBank references to "man, "homo sapeins", and "human" etc. as well as misspellings of these words are all converted to a standard term such as "human" (col. 27, lines 55-59).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to combine the teachings of the cited references because **Seilhamer's** teaching would have allowed **Kane's** to provide more accurate search results by standardizes the data from

various sources to conform with the pre-determined nomenclature as suggested in (col. 27, lines 14-23; col. 27, lines 61-62).

Regarding claim 12, **Kane** further teaches wherein **'the converter separates, rearranges, or combines the fields and the data items form the first records, into the uniform format'** as the source data is broken down into fields formatted to match the format of corresponding fields of the target database (col. 7, lines 17-23).

Regarding claim 13, **Kane** further teaches wherein **'before the normalizer converts the balance of the one or more data items, it separates the one or more data items into components'** as it is necessary to extract data from each source database so that the extracted data will be broken down into fields formatted to match the format of corresponding fields of the target database (col. 7, lines 17-23).

Regarding claim 14, **Kane** further teaches wherein **'the data masher determines that the first selected record and the second selected record are equivalent if the value is above a predetermined threshold level value'** as the scoring criteria is applied to determine minimum threshold for a match and which record has the highest ranking for a match. Data in the target database always being kept up to date and only source records that have a higher accuracy ranking than the fields in the target database are updated (col. 9, lines 18-58 and abstract; col. 12, lines 60-67).

Regarding claim 15, **Kane** further teaches software recordable in a tangible medium which includes machine readable instructions from performing a process for building a database, which stores records corresponding to a plurality of data items, the process comprising:

a). **'receiving at least a collection of first data items and a collection of second data items'** as processing source databases because record layouts between various source databases 101, 102, and 103 and the target database will not be identical and it is likely that different databases contain information about the same audience member, but that data within the databases may be inconsistent (col. 7, lines 17-23; Fig. 1; col. 4, lines 30-49);

b). **'disposing the first data items in a plurality of fields arranged in a predetermined format to form a first assemblage'** as it is necessary to extract data from each source database so that the extracted data will be broken down into fields formatted to match the format of corresponding fields of the target database (col. 7, lines 17-23);

c). **'disposing the second data items in a plurality of fields arranged in the predetermined format to form a second assemblage'** as it is necessary to extract data from each source database so that the extracted data will be broken down into fields formatted to match the format of corresponding fields of the target database Moreover, it may be desirable to use multiple source databases to update the target database (col. 7, lines 17-23; col. 8, lines 23-35);

e). **'maintaining a record, in the database, having a plurality of fields arranged in the predetermined format'** as the target database 108 of audience member records may have any desired record layout, but the indicated layout is suitable (col. 3, line 50 – col. 4, line 14); and

f). **'determining whether a first particular data item in the predetermined nomenclature in a selected field of the first assemblage is identical to a second particular data item in the predetermined nomenclature in a field of the second assemblage corresponding to the selected field'** as all source database records are first extracted, then, on a record by record basis, a matching record in the target database is identified and the data for the matched record in the target database is updated (Fig. 3 and col. 13, lines 47-52)(col. 9, lines 24-52). A representative criteria for determining whether a match exists between a source and a target record - **a negative number is a miss**, 0 is used when no data for a field is source, and **a 100 represents a perfect match** (i.e., identical) (col. 10, lines 18-49).

g). **'if it is determined that the first particular data item in the predetermined nomenclature is identical to the second particular data item in the predetermined nomenclature, including in a field in the record a selected one of (1) the first data items in the predetermined nomenclature and (2) the second data items in the predetermined nomenclature'** as all source database records are first extracted, then, on a record by record basis, a matching record in the target database is identified and the data for the matched record in the target database is updated (Fig. 3 and col. 13, lines 47-52)(col. 9, lines 24-52). For each field

group in the new target database record, the source database is specified for the fields name datasource_id, address_datasource_id, phone_datasource_id, and email_datasource_id, to indicate that the new target record information came from the particular source database (col. 12, lines 20-27).

d). **Kane** does not explicitly teach modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature.

Seilhamer, however, teaches **'modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature'** as a species filter may be employed to standardize all references to a particular species. For example, GenBank references to "man, "homo sapeins", and "human" etc. as well as misspellings of these words are all converted to a standard term such as "human" (col. 27, lines 55-59).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to combine the teachings of the cited references because **Seilhamer's** teaching would have allowed **Kane's** to provide more accurate search results by standardizes the data from various sources to conform with the pre-determined nomenclature as suggested in (col. 27, lines 14-23; col. 27, lines 61-62).

Regarding claim 16, **Kane** further teaches **'wherein at (b) and (c) the first data items and the second data items are separated, rearranged, or combined to form the first assemblage and the second assemblage respectively'** as the source data is broken down into fields formatted to match the format of corresponding fields of the target database (col. 7, lines 17-23).

Regarding claim 17, **Kane** further teaches **'wherein at (f) the first data items and the second data items are assigned accuracy rankings and are selected based upon their accuracy rankings'** as the scoring criteria is applied to determine minimum threshold for a match and which record has the highest ranking for a match. Data in the target database always being kept up to date and only source records that have a higher accuracy ranking than the fields in the target database are updated (col. 9, lines 18-58 and abstract; col. 12, lines 60-67).

Regarding claim 18, **Kane** further teaches a method for maintaining records in a database comprising:

a). **'converting first records containing data items to uniform records having the data items organized in a uniform format, wherein the uniform records have one or more of the data items conforming to a predetermined nomenclature and a balance of the one or more data items not conforming to the predetermined nomenclature'** as record layouts between various source databases 101, 102, and 103 and the target database will not be identical and it is likely that different databases contain information about the same audience member, but that data within the databases may be inconsistent. It is necessary to extract data from *each* source database so that the extracted data will be broken down into fields formatted to match the format of corresponding fields of the target database and compare the source records and update the target database with selected source records that have a higher accuracy ranking (col. 7, lines 17-23; Fig. 1; col. 4, lines 30-49; col. 9, lines 18-57);

'selecting, from among the second record, at least a first selected record and a second selected record based on a value representing a number of corresponding fields in the first selected record and the second selected record having identical data items therein' as comparing the source records and update the target database with selected source records that have a higher accuracy ranking (col. 9, lines 18-57; col. 4, lines 30-50). as all source database records are first extracted, then, on a record by record basis, a matching record in the target database is identified and the data for the matched record in the target database is updated (Fig. 3 and col. 13, lines 47-52)(col. 9, lines 24-52). A representative criteria for determining whether a match exists between a source and a target record - a negative number is a miss, 0 is used when no data for a field is source, and a 100 represents a perfect match (i.e., identical) (col. 10, lines 18-49).

c). **'selecting data items from the first selected record and second selected recorded to form a third data record, the third data record being stored in the database'** as comparing the source records and update the target database with selected source records that have a higher accuracy ranking (col. 9, lines 18-57; col. 4, lines 30-50).

b). **Kane** does not explicitly teach converting the balance of the one or more data items to the predetermined nomenclature, producing a collection of second records each having the one or more of the data items that conform to a predetermined nomenclature and the balance of the one or more data items that are in the predetermined nomenclature.

Seilhamer, however, teaches **'converting the balance of the one or more data items to the predetermined nomenclature, producing a collection of second records each having the one or more of the data items that conform to a predetermined nomenclature and the balance of the one or more data items that are in the predetermined nomenclature'** as a species filter may be employed to standardize all references to a particular species. For example, GenBank references to "man, "homo sapeins", and "human" etc. as well as misspellings of these words are all converted to a standard term such as "human" (col. 27, lines 55-59).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to combine the teachings of the cited references because **Seilhamer's** teaching would have allowed **Kane's** to provide more accurate search results by standardizes the data from various sources to conform with the pre-determined nomenclature as suggested in (col. 27, lines 14-23; col. 27, lines 61-62).

Regarding claim 19, **Kane** further teaches **'wherein the fields and the data items, from the first records, are separated, rearranged, or combined'** as the source data is broken down into fields formatted to match the format of corresponding fields of the target database (col. 7, lines 17-23).

Regarding claim 20, **Kane** further teaches wherein **'the balance of the one or more data items are separated into components'** as the source data is broken down into fields formatted to match the format of corresponding fields of the target database (col. 7, lines 17-23).

Kane does not explicitly teach converting the balance of the one or more data items into the *predetermined nomenclature*.

Seilhamer, however, teaches **'converting the balance of the one or more data items into the predetermined nomenclature'** as a species filter may be employed to standardize all references to a particular species (col. 27, lines 55-59).

Regarding claim 21, **Kane** further teaches wherein **'the first selected record and the second selected record are determined to be equivalent if the value is above a predetermined threshold level value'** as the scoring criteria is applied to determine minimum threshold for a match and which record has the highest ranking for a match. Data in the target database always being kept up to date and only source records that have a higher accuracy ranking than the fields in the target database are updated (col. 9, lines 18-58 and abstract; col. 12, lines 60-67).

Examiner's Remarks

5. Examiner would like to thanks Mr. Yip for his time in conducting interviews with the Examiner on dates 16 and 17 February 2005. The interviews help Examiner to better understand Applicant's claimed limitations and facilitate advancing the prosecution of the present application.

First, Applicant argues that nowhere do Kane and Seilhamer, individually or in combination, teach or suggest *selecting data items from (a) first assemblage or (b) a second assemblage when a "particular" data item (e.g., a telephone number) in the first assemblage is "identical" to a corresponding data item in the second assemblage*, as amended claim 1 now recites.

In response to the preceding arguments, Examiner respectfully submits that Kane teaches the above mentioned limitation as *all source database records are first extracted, then, on a record by record basis, a matching record in the target database is identified and the data for the matched record in the target database is updated* (Fig. 3 and col. 13, lines 47-52)(col. 9, lines 24-52). Applicant's first assemblage is equivalent with Kane's one of the source databases (101, 102, 103) (col. 11, lines 25-27). One of the source databases also may be the target database 108 itself (col. 12, lines 49-51). If the data items of the source database 101 is equivalent to Applicant's first assemblage and the data items of the source database 2 102 is equivalent to Applicant's second assemblage. Kane further teaches representative criteria for determining whether a match exists between a source and a target record - **a negative number is a miss**, 0 is used when no data for a field is source, and **a 100 represents a perfect match** (col. 10, lines 18-49). Suppose that a specific record from each of Source Database 2 102 and Source Database 3 103 have been found to match a specific record of the Target database 108, only those fields that have higher ranking than the fields in the target database are updated (col. 9, lines 24-52 and abstract). When Source Database 2 is processed, the email field of the Target Database will be updated, because rank of the field email_id of Source Database 2 (4) outranks the rank of the

email field of Source Database 1 (7), which is the source of the current content of the email field in the Target Database (col. 9,lines 32-37).

Based on the above, Examiner submits that Kane teaches the newly recite claimed limitation: "...*selecting data items from (a) first assemblage or (b) a second assemblage when a "particular" data item (e.g., a telephone number) in the first assemblage is "identical" to a corresponding data item in the second assemblage*" .

Second, Applicant further argues that Kane teaches away from the invention by using a sliding window" technique comparing every character in each word in a source record with every character in each word of a target record (col. 11, lines 1-5), and by determining whether or not to copy data only after all the characters and words in both records have been examined. By contrast, in accordance with the invention represented by amended claim 1, the determination of the above selection of data items from (a) or (b) is made as soon as the particular data item in the first assemblage and the corresponding data item in the second assemblages are found identical.

In response to the preceding arguments, Examiner respectfully submits that Kane teaches for each record of the set of target audience member record, comparing the field to the corresponding field of the source audience member record and assigning the plurality of match scores for the target audience member record for the field based on the extent of the match (col. 3, lines 18-27). Kane further teaches representative criteria for determining whether a match exists between a source and a target record - **a negative number is a miss**, 0 is used when no

data for a field is source, and **a 100 represents a perfect match** (i.e., identical) (col. 10, lines 18-49).

Based on the above, Examiner submits that Kane teaches the limitation "...the determination of the above selection of data items from (a) or (b) is made as soon as the particular data item in the first assemblage and the corresponding data item in the second assemblages are found identical" as claimed. Further, Applicant is reminded that in order to disqualify Kane based on a "teach away" reasoning, the reference has to *explicitly* suggest or disclose the so-called teach away steps: "Kane has to disclose the step of NOT or CAN NOT use the **fields comparison** between the source and target records.

Third, Applicant argues that Seilhamer also fails to teach or suggest the features "...*selecting data items from (a) first assemblage or (b) a second assemblage when a "particular" data item (e.g., a telephone number) in the first assemblage is "identical" to a corresponding data item in the second assemblage*".

In response to the preceding arguments, Examiner respectfully submits that Kane teaches the above features as spelled out in details from the above paragraphs. Therefore, Seilhamer does not necessary has to teach the same limitation. Seilhamer was brought in to supplement the missing feature of "*modifying selected ones of the first data items and the second data items to conform to predetermined nomenclature*".

Fourth, Applicant argues that Kane and Seilhamer, individually or in combination, teach or suggest, fail to disclose or suggest "determining a [score] value representing a number of corresponding fields in the first assemblage and the second assemblage having identical data items" and "data masher based on the [score] value representing a number of corresponding fields in the first assemblage and the second assemblage having identical data items" as amended claim 7 now recites.

In response to the preceding arguments, Examiner respectfully submits that Kane teaches the above mentioned limitation as representative criteria for determining whether a match exists between a source and a target record - **a negative number is a miss**, 0 is used when no data for a field is source, and **a 100 represents a perfect match.** Certain fields appearing in both the source and target databases may have weights assigned to them, such as: FNAME_WEIGHT=7, LNAME_WEIGHT=8, and PHONE_WEIGHT=5, etc... Each of these is a multiplier of the match score of –100 to 100 (col. 10, lines 18-49).

Based on the above, Examiner submits that Kane teaches the claimed limitations: "determining a [score] value representing a number of corresponding fields in the first assemblage and the second assemblage having identical data items" and "data masher based on the [score] value representing a number of corresponding fields in the first assemblage and the second assemblage having identical data items" as amended claims 7 and 11 now recites.

Conclusion

5. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Crozier (US 5701423 A)

Smith et al. (US 6052693 A)

Papierniak et al. (US 6151601 A)

Thompson et al. (US 6668253 B1)

Wheeler et al. (US 20020055932 A1)

Johnson (PCT/US00/01084)

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Leslie Wong whose telephone number is (571) 272-4120. The examiner can normally be reached on Monday to Friday 9:30am - 6:30 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John E Breene can be reached on (571) 272-4107. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Leslie Wong
Patent Examiner
Art Unit 2167

LW
February 25, 2005